INTERNATIONAL JOURNAL OF INNOVATIVE RESEARCH

IN COMPUTER & COMMUNICATION ENGINEERING

INTERNATIONAL STANDARD SERIAL NUMBER INDIA

Impact Factor: 8.379

Audio to Sign Language Translator using Artificial Neural Network Algorithms

Dr. D.J Samatha Naidu*[1], M. Bhanu Vijaya Venkata Krishna*[2]

*[1,] Principal, Annamacharya PG College of Computer Studies, Rajampet, Andhra Pradesh, India

*[2,] Student MCA, Annamacharya PG College of Computer Studies, Rajampet, Andhra Pradesh, India

ABSTRACT: Deaf people always miss out the fun that a normal person does,be it communication, playing computer games, participating in seminars or video conferences, etc. Communication is the most common difficulty in normal people, and not every normal person knows sign language. The aim of the project is to develop a communication system for the deaf. It converts the voice message into sign language. This system takes audio input, converts that voice-recorded message into text, and displays meaningful predefined Indian Sign Language images or GIFs. By using this system, communication between normal people and deaf people becomes easier.

KEYWORDS: Sign Language, Audio message, GIFs, Deaf People, Communication.

I. INTRODUCTION

The Sign language is the mother tongue of the deaf. It involves a combination of hand movements, hand or body and facial expressions. There are 135 types of sign languages worldwide. They are some Sign Languages is American Sign Language (ASL), Indian Sign Language (ISL), British Sign Language (BSL), Australian Sign Language (Auslan) and many more. I am using Indian Sign Language in the project. Through this system, deaf communities can enjoy all kinds of things that normal people do, from daily communication to accessing information. Sign language is a communication language that deaf people use with their face, hands or eyes using the speech channel.

The sign language recognition is used to recognize sign language for deaf and mute people. Gesture recognition is an important topic because segmentation of a foreground object from a cluttered background is a difficult problem. There is a difference between a person looking at an image and a computer looking at an image. It is easier for people to understand what is in the picture, but not for the computer. This is because computer vision problems remain a challenge. Sign language is a language consisting of gestures, facial expressions and body postures made with hands and other movements, mainly used by deaf or hard of hearing people to easily express their thoughts or communicate with other people. . Sign language is very important for deaf people who care about their emotional, social and linguistic growth. The first language of the deaf is sign language, which is progressing bilingually in the teaching of both the national sign language and the national written or spoken language. There are different deaf communities in different parts of the world, therefore the sign language of these communities is different. The different sign languages used in different communities are American Sign Language in America, British Sign Language in Great Britain, Indian Sign Language in India, etc. To express the thoughts and communicate with each other.

II. RELATED WORK

In [2] authors As per Amit Kumar Shinde on his study of sign language to text and vice versa in Marathi Sign language recognition is one of the most important types of research and it is the most natural and common way of communication for the people with hearing problems. A hand gesture recognition system can help deaf persons to communicate with normal people in the absence of an interpreter. The system works both in offline mode and through web camera.

Neha Poddar, Shrushti Rao, Shruti Sawant, Vrushali Seamanship, Prof. Sumita Chandak in their paper discussed about the prevalence of deafness in India is fairly significant as it is the second most common cause of disability. A portable interpreting device which converts higher mathematics sign language into corresponding text and voice can be very useful for the deaf people and solve many difficulties.

The glove based deaf-mute communication interpreter introduced by Anbarasi Rajamohan, Hemavathy R., Dhanalakshmi is great research. The glove comprises of five flex sensors, tactile sensors and accelerometer. The controller matches the gesture with pre-stored outputs. The evaluation of interpreter was carried out for ten letters _A,, _B,, _C,, _D,, _F,, _I,, _L,, _O,, _M,, _N,, _T,, _S,, _W,,.

III. PROPOSED ALGORITHM

Convolutional Neural Network

Convolutional Neural Network is one of the main categories to do image classification and image recognition in neural networks. Scene labelling, objects detections, and face recognition, etc., are some of the areas where convolutional neural networks are widely used.

CNN takes an image as input, which is classified and process under a certain category such as dog, cat, lion, tiger, etc. The computer sees an image as an array of pixels and depends on the resolution of the image. Based on image resolution, it will see as $h * w * d$, where h=height w=width and d= dimension. For example, An RGB image is $6 * 6 * 3$ array of the matrix, and the grayscale image is $4 * 4 * 1$ array of the matrix.

In CNN, each input image will pass through a sequence of convolution layers along with pooling, fully connected layers, filters (Also known as kernels). After that, we will apply the Soft-max function to classify an object with probabilistic values 0 and 1.
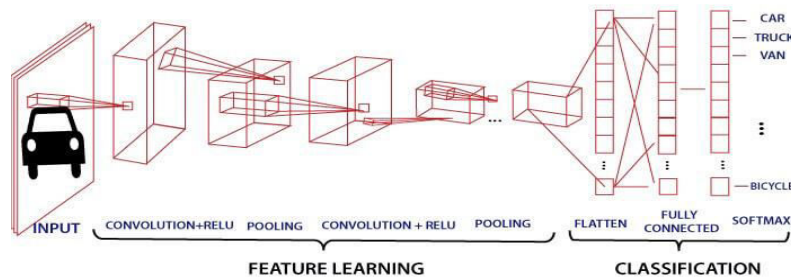

FIG. RELATED CNN ALGORITHM

Convolution Layer

Convolution layer is the first layer to extract features from an input image. By learning image features using a small square of input data, the convolutional layer preserves the relationship between pixels. It is a mathematical operation which takes two inputs such as image matrix and a kernel or filter.

Strides

Stride is the number of pixels which are shift over the input matrix. When the stride is equalled to 1, then we move the filters to 1 pixel at a time and similarly, if the stride is equalled to 2, then we move the filters to 2 pixels at a time. The following figure shows that the convolution would work with a stride of 2.

Padding

Padding plays a crucial role in building the convolutional neural network. If the image will get shrink and if we will take a neural network with 100's of layers on it, it will give us a small image after filtered in the end.
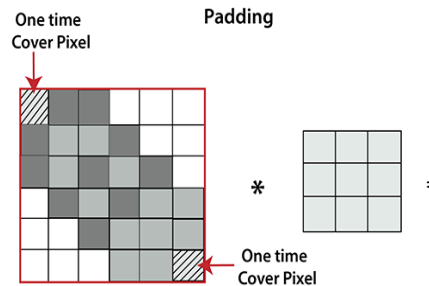
It is clear from the above picture that the pixel in the corner will only get covers one time, but the middle pixel will get covered more than once. It means that we have more information on that middle pixel, so there are two downsides:

- Shrinking outputs
- Losing information on the corner of the image.

Pooling Layer
Pooling layer plays an important role in pre-processing of an image. Pooling layer reduces the number of parameters when the images are too large. Pooling is "downscaling" of the image obtained from the previous layers. It can be compared to shrinking an image to reduce its pixel density. Spatial pooling is also called down sampling or subsampling,

Max Pooling
Max pooling is a sample-based discretization process. Its main objective is to downscale an input representation, reducing its dimensionality and allowing for the assumption to be made about features contained in the sub-region binned.
Max pooling is done by applying a max filter to non-overlapping sub-regions of the initial representation.
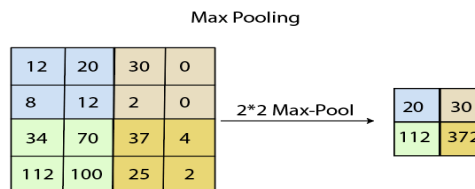
IV. PSEUDO CODE

Step 1: Open Application.
Step 2: Signup or login.
Step 3: Input the click on microphone to speak or Input enter text.
Step 4: Click on submit.
Step 5: Input is Process by the system.
Step 6: Click on play/pause for display of animation.
Step 7: Show the Required result.
Step 8: Close

V. SIMULATION RESULTS

The proposed is to develop a communication system from the deaf people it's convert the audio message into the sign language. Where system takes audio as input convert this audio record message into text and display the relevant data Indian sign language. Images or GIFS. Which are predefined. It is useful for the for both normal and deaf people. Fig 1. User open screen you can see audio to sign language translator home page. Fig 2. User sign up can fill the user name, password and password conformation and click on sing up. Fig. 3 User can here display the input audio taken to the

user. Its display the screen.Fig.4 User click on submit screen represent audio is convert key words in sentence. Fig 5. Next User click on play/pause is using sign represents. Fig 6. Finally display the sign images to the user is showing.
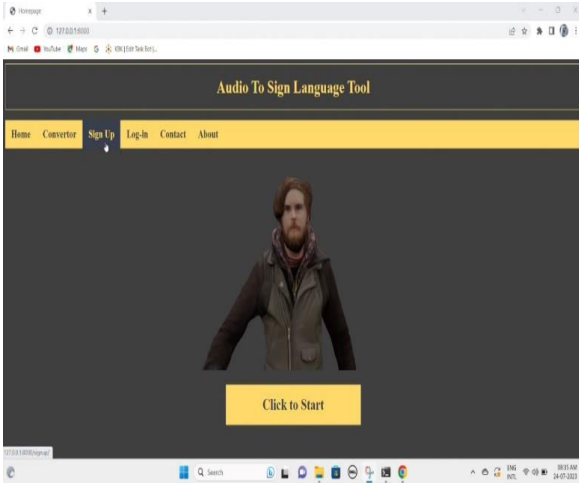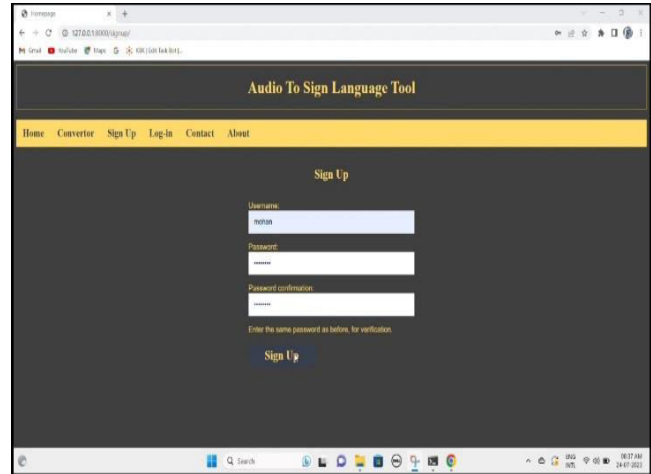

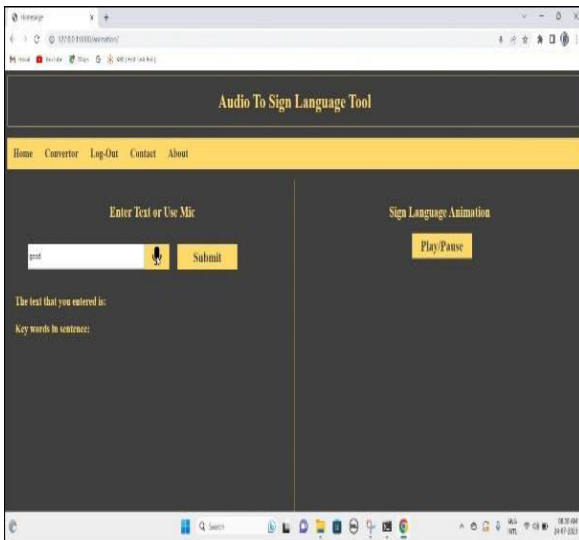Fig.1.Home Page
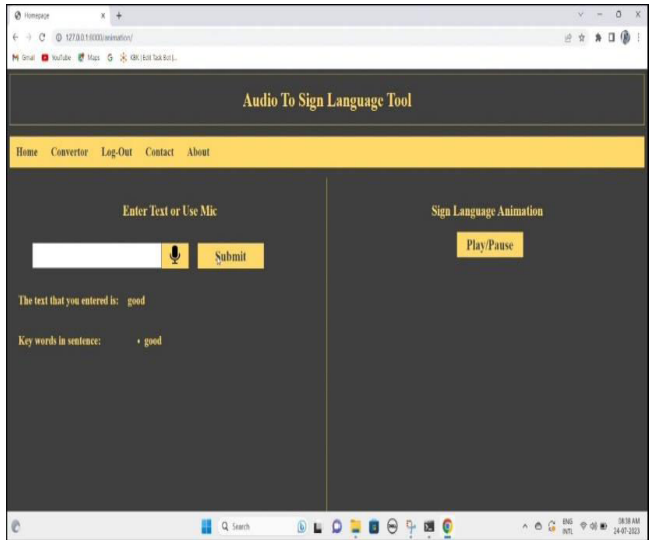

Fig. 2. Sign up page

Fig. 3. Input speech or Audio Screen
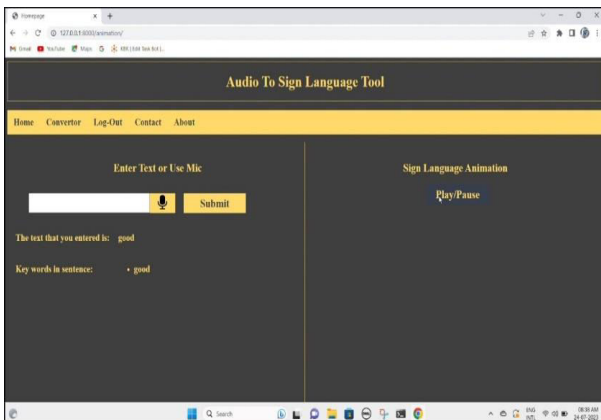

Fig 4. Click on the submit Screen

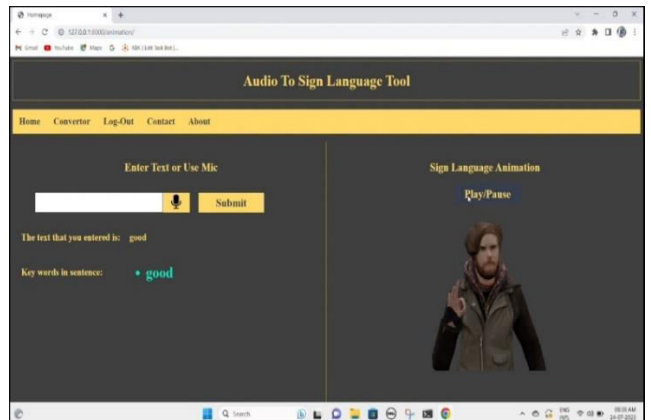Fig. 5. Click on the play/Pause Screen

Fig.6. Display Sign Image Screen

VI. CONCLUSION AND FUTURE WORK

Finally, I conclude, the Sign language translator is one of the useful tools is easy to the communication between the deaf and mute communities and normal society. Though sign language can be implemented to communicate, the target person must have an idea of the sign language which is not possible always. This was meant to be a prototype to check the feasibility of recognizing sign language. The normal people can communities with deaf or dumb using sign language and the audio to text will be converted to sign images.

REFERENCES

1. B. R. Callejas Bedregal, G. P. Dimuro and A. C. Rocha Costa," Interval fuzzy rule-based hand gesture recognition," in 12th GAMM-IMACS International Symposium on Scientific Computing, Computer Arithmetic and Validated Num erics, SCAN 2006, September 26, 2006 - September 29, 2006, Duisburg, United states, 2006.
2. M. G. Ceruti, V. V. Dinh, N. X. Tran, H. Van Phan, L. T. Duffy, T. Ton, G. Leonard, E. Medina, O. Amezcua, S. Fugate, G. J. Rogers, R. Luna, and J. Ellen," Wireless communication glove apparatus for motion tracking, gesture recognition, data transmission, and reception in extreme environments," in 24th Annual ACM Symposium on Applied Computing, SAC 2009, March 8, 2009 - March 12, 2009, Honolulu, HI, United states, 2009, pp. 172-176.
3. C. Oz and M. C. Leu," Human-computer interaction system with artificial neural network using motion tracker and data glove," in 1st International Conference on Pattern Recognition and Machine Intelligence, PReMI 2005, December 20, 2005 December 22, 2005, Kolkata, India, 2005, pp. 280-286.
4. M. Kato, Y. Chen and G. Xu," Articulated hand tracking by PCA-ICA approach," in FGR 2006: 7th International Conference on Automatic Face and Gesture Recognition, April 10, 2006 - April 12, 2006, Southampton, United Kingdom, 2006, pp. 329-334.
5. H. Guan, R. S. Feris and M. Turk," The Isometric Self Organizing Map for 3D hand pose estimation," in FGR 2006: 7th International Conference on Automatic Face and Gesture Recognition, April 10, 2006 - April 12, 2006, Southampton, United Kingdom, 2006, pp. 263-268.
6. M. Bicego, U. Castellani and V. Murino," A hidden Markov model approach for appearance-based 3D object recognition," Pattern Recognition Letters, vol. 26, pp. 2588-2599, 2005-1201 2005.
7. Y. Cui and J. Weng," Appearance-based hand sign recognition from intensity image sequences," Computer Vision and Image Understanding, vol. 78, pp. 157-176, 2000.
8. P. Kakumanu, S. Makrogiannis and N. Bourbakis," A survey of skin-color modeling and detection methods," Pattern Recognition, vol. 40, pp. 1106-1122, 2007.

INTERNATIONAL JOURNAL OF INNOVATIVE RESEARCH

IN COMPUTER & COMMUNICATION ENGINEERING

📱 9940 572 462 🟢 6381 907 438 ✉ ijircce@gmail.com

Scan to save the contact details